

WO 2005/093573 A2

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
6. Oktober 2005 (06.10.2005)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2005/093573 A2

(51) Internationale Patentklassifikation⁷: **G06F 11/00**

[AT/AT]; Schwarzgrubergasse 16, Top 3, A-1100 Wien (AT).

(21) Internationales Aktenzeichen: PCT/EP2005/050980

(74) Anwalt: BERG, Peter; Siemens Aktiengesellschaft, Postfach 22 16 34, 80506 München (DE).

(22) Internationales Anmeldedatum:

4. März 2005 (04.03.2005)

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Einreichungssprache:

Deutsch

(26) Veröffentlichungssprache:

Deutsch

(30) Angaben zur Priorität:

10 2004 014 195.9 23. März 2004 (23.03.2004) DE

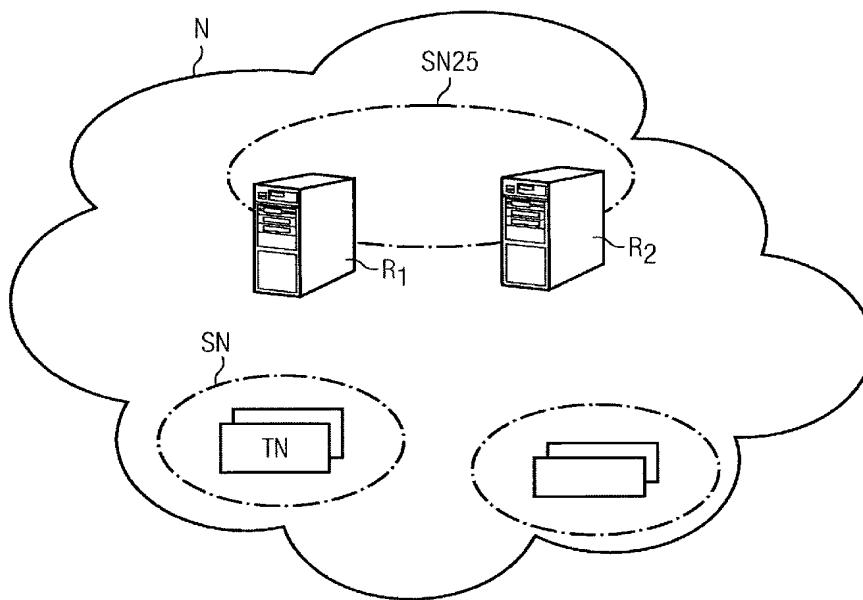
(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): SIEMENS AKTIENGESELLSCHAFT [DE/DE]; Wittelsbacherplatz 2, 80333 München (DE).

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR REDUNDANT DATA MANAGEMENT IN COMPUTER NETWORKS

(54) Bezeichnung: VERFAHREN ZUR REDUNDANTEN DATENHALTUNG IN COMPUTERNETZWERKEN



(57) Abstract: The invention relates to a method for redundant data management in computer networks, wherein data is consistently stored in at least two computers (R1, R2). The invention is characterized in that the data management application programs used for the at least two computers have a coinciding address identification, which said application programs have as component of a virtual sub-network (SN 25) of a computer network (N). The computers are independent of one another and have routing functions by means of which the application programs are accessed. This makes it possible to implement redundant data management in a simple manner that is invisible for the user.

[Fortsetzung auf der nächsten Seite]



GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

Veröffentlicht:

- *ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts*

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur redundanten Datenhaltung in Computernetzwerken, bei dem Daten auf zumindest zwei Rechnern (R1, R2) konsistent gespeichert werden, dadurch gekennzeichnet, dass die für die Datenhaltung vorgesehenen Anwendungsprogramme der zumindest zwei Rechner eine übereinstimmende Adresskennung aufweisen, welche diese Anwendungsprogramme als Bestandteil eines virtuellen Subnetzes (SN 25) eines Computernetzwerkes (N) aufweist und dass die voneinander unabhängigen Rechner Routingfunktionen aufweisen über welche der Zugriff auf die Anwendungsprogramme erfolgt. Damit wird auf einfache und für den Nutzer unsichtbare Weise eine redundante Datenhaltung ermöglicht.

Beschreibung**Verfahren zur redundanten Datenhaltung in Computernetzwerken**

5

Technisches Gebiet

Die Erfindung betrifft ein Verfahren zur redundanten Datenhaltung in einem Computernetzwerk, bei dem Daten auf zumindest zwei Rechnern eines Netzwerkes konsistent gespeichert werden.

10

Stand der Technik

Computernetzwerke haben sich in den letzten Jahren zu einem bedeutenden Kommunikationsmedium entwickelt, über welches eine Vielzahl von Diensten angeboten wird. Das weltweit bekannteste Computernetzwerk ist das Internet, wie die Gesamtheit aller miteinander verbundenen und das Internet Protokoll IP als Transportprotokoll nutzenden Netze bezeichnet wird.

20

Das Internet Protokoll IP ist auf Ebene 3 des OSI-Schichtenmodells für den verbindungslosen Transport von Daten von einem Sender – gegebenenfalls über mehrere Netze – zum Empfänger zuständig, wobei weder Fehlererkennung noch Fehlerkorrektur erfolgt.

25

Auf dem Internet Protokoll setzen Transport Protokolle wie das Transmission Control Protocol TCP oder User Datagram Protocol UDP auf.

30

Das Transmission Control Protocol TCP ist ein verbindungsorientiertes Transportprotokoll, welches eine logische, sichere Voll duplex Punkt-zu-Punkt Verbindung ermöglicht. Es stellt dabei sicher, dass Datenpakete fehlerfrei und in der gewünschten Reihenfolge über ein

35

darunter liegendes Netz nach dem Internet Protokoll IP übertragen werden.

Das User Datagram Protocol UDP ermöglicht eine
5 verbindungslose Punkt-zu-Punkt Kommunikation. Die einzelnen Datenpakete (Datagramme) sind dabei voneinander unabhängig. Die Übertragung ist nicht gesichert. Dies bedeutet, dass Datenpakete verloren gehen können und dass die Empfangsreihenfolge nicht mit der Sendereihenfolge
10 übereinstimmen muss. Eine allfällige Sicherung des Datentransports hat, falls es erforderlich ist, auf der Applikationsebene zu erfolgen. UDP wird im Allgemeinen für einfache Dienste verwendet, bei denen auf eine einfache Anfrage (ein Datagramm) eine einfache Antwort (ein Datagramm)
15 erwartet wird.

Eine der beliebtesten Anwendungen des Internets ist das World Wide Web WWW, das mittels Hyperlinks den einfachen Zugriff auf multimediale Daten entfernter WWW-Server ermöglicht.
20 Betrachtet und dargestellt werden die Daten, die Texte, Audiosequenzen, Standbilder und Filme umfassen können, mittels WWW-Browser, einer lokalen Frontend Software. Das World Wide Web beruht auf dem Client/Server Prinzip, auf dem alle wichtigen Anwendungen im Internet, neben WWW auch
25 FTP oder News basieren.

Nach dem Client-Server Prinzip gestaltete Anwendungen verwenden auf der Benutzerseite ein Clientprogramm, das mit einem bestimmten Diensttrechner im Netz - dem Server - Daten
30 austauscht. Der Server ist dabei in der Regel für die Datenhaltung zuständig, während der Client die Präsentation dieser Daten und die Interaktion mit dem Benutzer übernimmt. Dazu bedienen sich Client und Server eines genau definierten Protokolls.

35 Betrieben werden die Server des Internets entweder von kommerziellen oder auch nicht kommerziellen Organisationen,

wie z.B. Universitäten, die verschiedene Online-Dienste, also Informationsdienste (Wetterbericht, Zugfahrpläne, Nachrichtendienste) und Kommunikationsdienste (E-Mail, Chat, Foren) anbieten, welche mittels entsprechenden
5 Anwendungsprogrammen realisiert werden.

Ein wesentliches Kriterium für die Qualität eines angebotenen Dienstes ist die Wahrscheinlichkeit, mit der ein potentieller Benutzer den Dienst auch nutzen kann, die entsprechende
10 Maßzahl - welche sich aus der Störzeit während einer bestimmten Beobachtungsdauer ergibt - die sogenannte Verfügbarkeit dieses Dienstes, hat maßgeblichen Einfluss auf die Zufriedenheit der Benutzer und damit auf den Erfolg des jeweiligen Dienstes.

15 Zur Erhöhung der Verfügbarkeit besteht daher oftmals die Forderung, Dienste bzw. die jeweils zugrundeliegenden Daten gegen Ausfälle der maßgeblichen Hardwarekomponenten des Computernetzes, also der Rechner und der Verbindungen
20 abzusichern. Dies geschieht üblicherweise durch redundante Datenhaltung auf mehreren voneinander unabhängigen Rechnern eines Netzwerkes. Dazu ist es beispielsweise aus dem Beitrag „FPGA-based load balancer for Internet Servers, Harik,L., Kayssi, A.“; erschienen in Microelectronics, The 14th
25 International Conference on 2002-ICM bekannt, mittels sogenanntem „Load Balancer“ eintreffende Nachrichten auf unterschiedliche Server zu verteilen. Der „Load Balancer“ bildet dabei einen „single point of failure“, d.h. bei Ausfall des „Load Balancers“ selbst ist keiner der
30 nachfolgenden Server erreichbar.

Darstellung der Erfindung

Der Erfindung liegt die Aufgabe zugrunde, ein verbessertes
35 Verfahren zur redundanten Datenhaltung in Computernetzen anzugeben.

Erfindungsgemäß geschieht dies mit einem Verfahren zur redundanten Datenhaltung in Computernetzwerken, bei dem Daten auf zumindest zwei Rechnern konsistent gespeichert werden,
5 wobei die für die Datenhaltung vorgesehenen Anwendungsprogramme der zumindest zwei Rechner eine übereinstimmende Adresskennung aufweisen, welche diese Anwendungsprogramme als Bestandteil eines virtuellen Subnetzes eines Computernetzwerkes aufweist und wobei die
10 voneinander unabhängigen Rechner Routingfunktionen aufweisen, über welche der Zugriff auf die Anwendungsprogramme erfolgt.

Die Erfindung ist besonders vorteilhaft bei Computernetzwerken anwendbar, bei denen die Kommunikation auf
15 dem Internet Protokoll beruht, und bei denen als Transportprotokolle das Transmission Control Protocol TCP oder das User Datagram Protocol UDP angewendet werden.

Günstig ist es weiterhin, wenn durch entsprechende
20 Gestaltung der Routingfunktion eine Zugriffsrangfolge auf die für die Datenhaltung vorgesehenen Anwendungsprogramme festgelegt wird.

Kurzbeschreibung der Zeichnung
25 Die Erfindung wird anhand einer Figur näher erläutert, welche in einer schematischen Darstellung die Adressierung in einem beispielhaften Computernetzwerk zeigt.

Ausführung der Erfindung
30 Die Adressierung im dargestellten Computernetzwerk N erfolgt nach dem Internet-Protokoll Version 4 durch die sogenannte IP(Internet-Protokoll)-Adresse, die jeden Netzeilnehmer auf Ebene 3 des OSI-Schichtenmodells, also auf einer logischen Ebene eindeutig kennzeichnet.

Die IP-Adresse ist nach dem Internet-Protokoll Version 4 ein 32-Bit-Wort, welches einen ersten Bereich aufweist, der zur Adressierung des Netzes herangezogen wird, und einen zweiten Bereich, mittels dem der Rechner im Netz identifiziert wird.

Die Größe der beiden Bereiche ist normiert, wir unterscheiden 5 Klassen von IP-Adressen:

- Klasse A Netze mit 7 bit Netzadressierung und 24 bit für die Teilnehmeradresse,
- Klasse B Netze mit 14 bit Netzadressierung und 16 bit Teilnehmeradressierung,
- Klasse C Netze mit 21 bit Netzadressierung und 8 bit Teilnehmeradresse.
- Klasse D für die Adressierung innerhalb des Netzes mit 28 bit Teilnehmeradressen und der für Forschungszwecke reservierten
- Klasse E

Die Klasse der IP-Adresse ist in den führenden bits dargestellt, Adressen der Klasse A sind dadurch erkenntlich, dass das führende bit den Wert 0 aufweist, Klasse B durch den Wert 10 der beiden führenden bits, Klasse C durch 110 der drei ersten bits usw.

Die Information über die Adressklasse benötigen die Router eines Netzwerkes um mit Hilfe der sogenannten Netzmasks aus der IP-Adresse eines Zielteilnehmers die Information über die Adresse des Zielnetzes herauszufiltern. Während die Netzadressen global verwaltet werden, erfolgt die Verwaltung der Teilnehmeradressen lokal durch den jeweiligen Netzbetreiber. Dieser hat dabei auch die Möglichkeit, durch entsprechende Vergabe der Teilnehmeradressen sein Netz durch Unterteilung in Subnetze zu strukturieren, ohne zusätzliche Netznummern zu verwenden. Dazu wird der Teilnehmeranteil der IP-Adresse in eine

- Subnetzwerkadresse und die

- eigentliche Teilnehmeradresse unterteilt.

Die IP-Adresse ist dann dreiteilig aufgebaut, sie umfasst die Netzwerkadresse, die Subnetzadresse und die eigentliche
5 Teilnehmeradresse.

Als einfaches Beispiel sei ein Netz der Adressklasse B angegeben, bei dem die IP-Adresse die Binärzahl 10 in den führenden beiden bits und weitere 14 bits für die Adresse des
10 Netzes aufweist. Der bei dieser Netzwerkklassse vorgesehene Teilnehmeradressierungsteil von 16 bit wird wiederum unterteilt in einen 8 bit Teil zur Adressierung der Subnetze SN1, ...SN 25, ..., SN 255 und in ein 8 bit Element zur Adressierung der Teilnehmer TN in den Subnetzen. Damit können
15 maximal 254 Subnetze mit jeweils 254 Teilnehmern adressiert werden.

20 In der allgemein üblichen 4 Oktett-Schreibweise der IP-Adressen, bei der die einzelnen Bytes (8 bit) durch dezimal numerische Werte dargestellt und durch Punkte getrennt werden, sieht eine typische IP-Adresse wie folgt aus:

172.16.115.54

25 Die Adresse des Netzes N selbst beträgt dabei 172.16.0.0. Diese umfasst bis zu 254 Subnetze SN1, ...SN 25, ..., SN 255 mit den Adressen

172.16.1.0, ..., 172.16.254.0.

30 Die 254 möglichen Teilnehmer TN in dem Subnetz SN 25 mit der Adresse 172.16.25.0 weisen dann folgende Adressen auf:
172.16.25.1 , ..., 172.16.25.254

35 Erfindungsgemäß werden nun auf zwei voneinander unabhängigen Rechnern R1, R2 des Netzes N die redundant zu haltenden Daten

konsistent gespeichert, wobei der Datenabgleich zwischen den redundanten Einheiten mit bekannten Verfahren sichergestellt wird. Dabei ist es möglich, dass die beiden Rechner R1, R2 einem Netz N eines Betreibers oder aber verschiedenen 5 miteinander verbundenen Netzen unterschiedlicher Betreiber angehören.

Den zugehörigen Anwendungsprogrammen der zwei Rechner wird eine übereinstimmende Adresskennung zugeordnet, welche diese 10 Anwendungsprogramme als Bestandteil eines virtuellen Subnetzes SN 25 eines Computernetzwerkes N ausweist, wie beispielsweise die Nummer 172.16.25.0. Das Anwendungsprogramm läuft aus Sicht des Benutzers auf einem Rechner in diesem 15 Subnetz SN 25, zum Beispiel auf einem Rechner mit der Adresse 172.16.25.1.

Dieses virtuelle Subnetz SN 25 ist über zwei Router logisch mit dem Gesamtnetz verbunden. Dazu sind die beiden voneinander unabhängigen Rechner R1, R2, welche unter anderem 20 die für die redundante und konsistente Speicherung der Daten vorgesehenen Anwendungsprogramme enthalten, als Router - ausgestaltet, über welche ausschließlich der Zugriff auf die Anwendungsprogramme - das virtuelle Subnetz - erfolgt.

25 Die übrigen Router der Netzwerke enthalten in ihren Routingtabellen daher einen Hinweis, dass Adressen des Typs 172.16.25.XXX, die ja das virtuelle Subnetz SN 25 betreffen, nur über einen der beiden genannten Router erreichbar sind. Der Eintrag in der Routingtabelle der übrigen Router enthält 30 in diesem Beispiel die Netzwerkadresse 172.16.25.0, die Netzwerkmaske 255.255.255.0, die Kosten für den Weg zu diesem Subnetz und das Interface, über das dieses Netz N erreichbar ist.

35 Alle Datenpakete mit diesen Zieladressen gelangen daher an einen der beiden Router und werden von diesem an das

entsprechende Anwendungsprogramm mit dem jeweils zugehörigen Speicherbereich weitergeleitet.

Wenn das Netz N, wie allgemein üblich mit einem dynamischen Routingverfahren ausgestattet ist, wird bei Belegung eines der beiden Anwendungsprogramme automatisch die nächste Anfrage zu dem jeweils günstigsten freien Speicherbereich geroutet. Sind die beiden Router, über die das virtuelle Subnetz SN 25 erreicht werden kann, über gleich teure Wege 10 erreichbar, so wird ein Router zufällig ausgewählt und damit eine Lastaufteilung vorgenommen.

Bei einem Netz mit einem statischen Routingverfahren muß eine 15 entsprechende Lastaufteilung bei der Festlegung der Routingtabellen erfolgen.

Das gegenständliche Ausführungsbeispiel beschreibt den Einsatz der Erfindung bei einem Computernetzwerk N mit einer 20 Adressierung nach IP-Protokoll Version 4, die Erfindung ist aber beispielsweise auch bei IP -Protokoll Version 6 vorteilhaft einsetzbar. Ebenso ist die Anwendung der Erfindung für alle gebräuchlichen auf IP aufsetzenden Transportprotokolle wie TCP oder UDP möglich.

Bei Anwendung des TCP-Protokolles muss allerdings 25 sichergestellt werden, dass der Zugriff jeweils über einen bestimmten Router erfolgt. Dies kann beispielsweise über die Kostenfunktionalität dynamischer Routingfunktionen erfolgen.

30

Patentansprüche

1. Verfahren zur redundanten Datenhaltung in Computernetzwerken, bei dem Daten auf zumindest zwei Rechnern konsistent gespeichert werden, dadurch gekennzeichnet, dass die für die Datenhaltung vorgesehenen Anwendungsprogramme der zumindest zwei Rechner (R1, R2) eine übereinstimmende Adresskennung aufweisen, welche diese Anwendungsprogramme als Bestandteil eines virtuellen Subnetzes (SN 25) eines Computernetzwerkes (N) aufweist und dass die voneinander unabhängigen Rechner Routingfunktionen aufweisen über welche der Zugriff auf die Anwendungsprogramme erfolgt.
- 15 2. Verfahren nach Anspruch 1 dadurch gekennzeichnet, dass die Kommunikation in den Computernetzwerken (N) auf dem Internet Protokoll beruht.
- 20 3. Verfahren nach Anspruch 2 dadurch gekennzeichnet, dass für die Kommunikation im Computernetzwerk überdies das Transmission Control Protocol genutzt wird.
- 25 4. Verfahren nach Anspruch 2 dadurch gekennzeichnet, dass für die Kommunikation im Computernetzwerk überdies das User Datagram Protocol genutzt wird.
- 30 5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass durch entsprechende Gestaltung der Routingfunktion eine Zugriffsrangfolge auf die für die Datenhaltung vorgesehenen Anwendungsprogramme festgelegt wird.

